数理解析研究所講究録
333巻 1978年 178-188

ランダムグラフの統計的解析

東北大学　工学部　　高見沢一秀
富士通株式会社　　滝内　政昭
東北大学　工学部　　西関　隆夫
　　　　　　　　　　斎藤　伸自

1．まえがき

近年，問題の規模が大きくなるにつれ，計算時間が指数関数的に増大するような解法しか知られていなかった組み合せ問題（彩色数問題などNP-完全とよばれているもの）に対して，実用的計算時間内で，近似最適解を求める近似算法がいくつか提案されている(1)(2)。これらの算法の理論的な有効性の評価ならびに比較は，極めて難しい問題であり，現実には，多数の例について，実際に計算を行う計算機実験に頼らざるをえない。このようなときに，ランダムなグラフデータ"ランダムグラフ"が必要となる。また，従来，グラフ問題の計算の複雑度は，その最悪値で評価されてきたが，実用的な観点からは，期待値評価の方が望ましい場合も多い。この計算

の複雑度の期待値評価を行う上でも，ランダムグラフを明確に定義し，その性質を明らかにすることが望まれる。

ランダムグラフは，いくつかの定義が与えられているが，その性質については知られていないことが多く，実際に計算機を用いてランダムグラフを発生させ，ランダムグラフ自身の統計的な解析を行った例もないようである。本文は，このような，ランダムグラフに関する統計的数値解析を試みたものである。なお，ランダムグラフの発生ならびにグラフ処理には，筆者らが開発したグラフ処理プログラム－GRAMP－[3]を用いている。

2. 準 備

本文では，グラフGを節点の集合Vと枝の集合Eの順序対と考え，$G=(V,E)$ と表わす。Eの各元が，2つの節点v，$w\in V$の順序対あるいは非順序対であるとき，Gをそれぞれ有向グラフあるいは無向グラフとよぶ。ここで，Gは自己閉路および同じ向きの並列枝を持たないものとする。

任意の2点間に道が存在する無向グラフを連結グラフとよぶ。極大な連結部分グラフを連結成分とよぶ。連結グラフが非可分であるというのは，どの3点v，w，aをとっても，vとwの間にaを通らない道が存在することである。連結グ

ラフは，極大な非可分部分グラフに一意に分割でき，それらを非可分成分とよぶ。有向グラフの各枝の向きを無視して得られる無向グラフが連結であるとき，元の有向グラフは弱連結であるという。また，任意の点から任意の点へゆく有向道が存在する有向グラフを強連結グラフとよぶ。弱連結グラフの極大な強連結部分グラフを強連結成分とよぶ。

3. ランダムグラフと既知の結果

いくつかのグラフがランダムグラフとして定義されているが，本文では文献(4)に従い，次のように定義する。

［定義 1］[(4)] ラベルの付けられた n 個の節点のいずれか2点を結ぶ総数 ${}_nC_2$ 本の枝の中から，m 本の枝をランダムに選んで得られるグラフを，n 個の節点および m 本の枝をもつランダムグラフとよぶ。（定義終）

有向ランダムグラフは，逆方向の枝も含めて，${}_nC_2 \times 2$ 本の枝から，ランダムに m 本選んで得られる有向グラフであると定義する。

ここで定義されたランダムグラフは，発生可能なラベル付きグラフをすべて等確率で発生させるものであり，ラベル無しグラフとして同形なものも区別して考えている。

Erdös，Rényi らは，上記のランダムグラフの連結度などに

ついて，いくつかの興味ある結果を示している[4][5]．以下に，その主なものを示す。

［定理1］[4]　節点数n，枝数mの間に

$$m=\left\lfloor \frac{n}{2}\log n + Cn \right\rfloor \qquad C:\ 実定数$$

の関係があり，nが十分大きいとする。このとき，ランダムグラフが連結である確率は

$$\exp(-e^{-2C}) \tag{1}$$

である。

［定理2］[5]　節点数n，枝数mの間に

$$m=\left\lfloor \frac{n}{2}\log n + Cn \right\rfloor \qquad C:\ 実定数$$

の関係があり，nが十分大きいとする。このとき，ランダムグラフが$k+1$個の連結成分を持ち，そのうちk個は孤立節点である確率は

$$\exp(-2kC-e^{-2C})/k! \tag{2}$$

である。

上の定理で，$\sum_{k=1}^{\infty}\exp(-2kC-e^{-2C})/k!=1$に注意すれば，ランダムグラフの連結成分はほとんど孤立節点であり，かつ大部分の節点は，1つの大きな連結成分に含まれていることがわかる。これは，本文で定義したランダムグラフの顕著な特徴の1つである。本文では，同様な特徴が，連結ランダムグラフの場合にもあることを実証する実験を行っている。

4. ランダムグラフの発生アルゴリズム

ランダムグラフ及び連結ランダムグラフの発生アルゴリズムを簡単に紹介する。

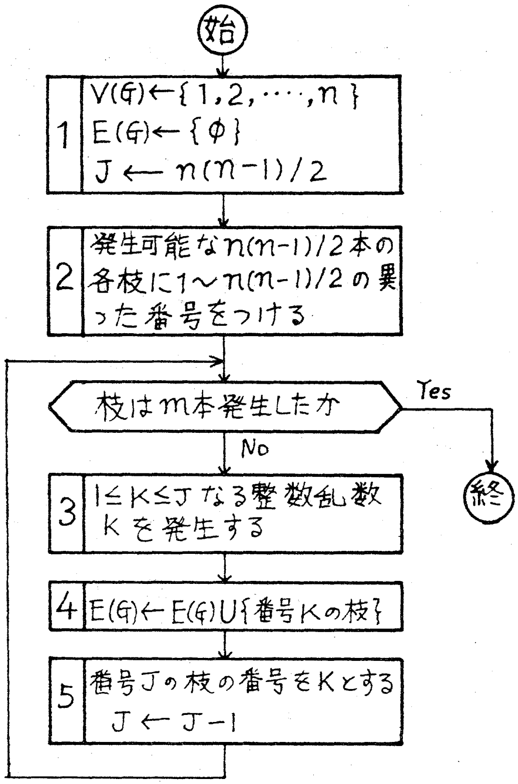

図1. ランダムグラフの発生アルゴリズム

ランダムグラフならびに連結ランダムグラフの発生プログラムはともに，発生可能な${}_nC_2$本の各枝に任意のラベル付けをすることを用いている。連結ランダムグラフは，ランダムグラフの中の連結なグラフだけを取り出すことによって生成することができるが，枝数が少ないグラフを生成するときには，この方法は効率が悪い。そこで，ここでは，始めに全ての節点を含む木を構成してから，残りの枝を付加して連結ランダムグラフを構成している。節点数n，枝数mのランダムグラフ及び連結ランダムグラフの生成アルゴリズムを図1，図2に示す．

図1のアルゴリズムのブロック1，2でかかる手数は$O(n^2)$である。ブロック3，4，5の手数は$O(1)$で，ループを回

る回数は m 回である。従って，ランダムグラフを発生する計算手数は $O(n^2)$ である。また，図3のアルゴリズムのブロック1から3までの手数は $O(n^2)$ である。ブロック4，7，8の手数は $O(1)$ であり，$T_1, T_2, \cdots, T_n$ に対して適当なデータ構造を考えれば，5及び6は，各々 $O(1)$, $O(\log n)$ の手数で実行できる。各ブロックの実行回数は，ブロック6が $n-1$ 回，7と8が m 回，4と5は期待値で n^2 回である。従って図2のアルゴリズムの手数は，期待値で $O(n^2)$ である。なお，両

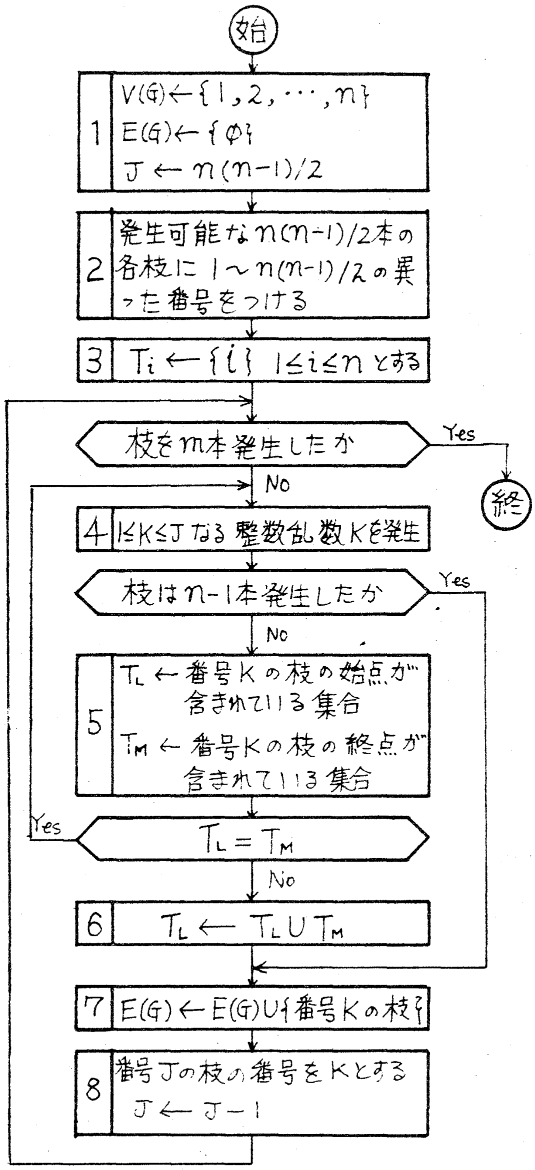

図2. 連結ランダムグラフの発生アルゴリズム

アルゴリズムとも，Shuffle の手法[7]を用いて，同一枝が重複して発生することがないようにしている（図1のブロック5，図2のブロック8）。

5. 実験結果

グラフ処理プログラムを用いて次のような実験を行った。

5.1 連結グラフの発生確率

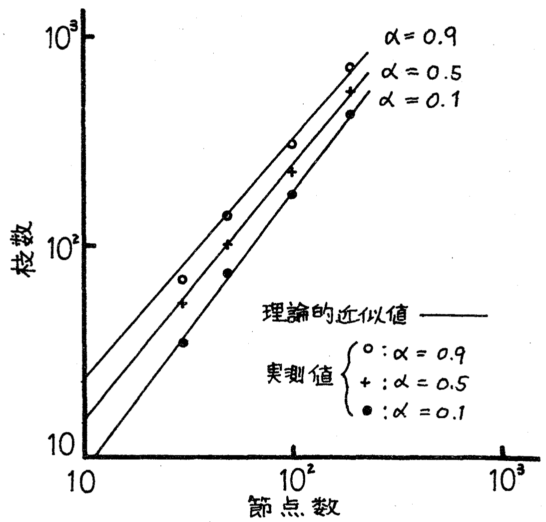

図3. 連結グラフの発生率αをパラメータとしたときの枝数と節点数の関係

ランダムグラフを多数発生させ，連結グラフが出現する確率を測定する。連結グラフの発生確率は(1)式で与えられるが，ここでは，この理論近似値と実測値の比較を行う。実験は，同じ節点数，枝数を持つグラフを100個ずつ発生し，その中に出現する連結グラフの個数を調べている。図3に，連結グラフの発生率が，それぞれ0.1，0.5，0.9となるときの節点数と枝数の関係を示す。実線が(1)式より計算される値である。定理1は，節点数が十分大きいという仮定をおいているが，節点数が100程度でも(1)式はかなり良い近似となっているといえよう。

5.2 ランダムグラフにおける divide-and-conquer 法の有効性

グラフ問題において，グラフ全体に関する問題が，グラフの個々の連結成分あるいは非可分成分に関する問題に帰着できる場合には，グラフをそのようないくつかの部分グラフに分割して，より小規模な問題に帰着させて解く場合がある(divide-and-conquer)。節点数nのグラフを，節点数がそれぞれ$n_1, n_2, \cdots, n_g$の部分グラフに分割できるとすれば，もともと

$O(n^2)$の計算時間を必要としたアルゴリズムは，この手法を用いて計算速度が $n^2/\sum_{i=1}^{q} n_i^2$ 倍になる。本文では，ランダムグラフに対するdivide-and-conquer 法の有効性を調べるために，ランダムグラフおよび連結ランダムグラフを，連結成分あるいは非可分成分に分割して，それぞれについて $n^2/\sum_{i=1}^{q} n_i^2$ を算出している。図4，5にその結果を示す。

実験結果より，枝数が節点数の倍以上のグラフでは，divide-and-conquer法の効果がほとんどないことがわかる。これは，ランダムグラフでは，節点の大部分が，1つの連結あるいは非可分成分に含まれてしまうことによるものと考えられる。このことは，次の実験で更に明確になる。

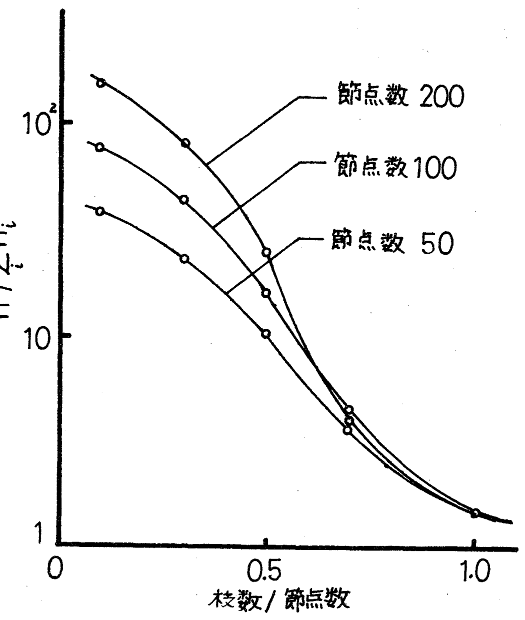

図4. 連結成分に分割したときの$n^2/\sum_i n_i^2$と枝数/節点数の関係

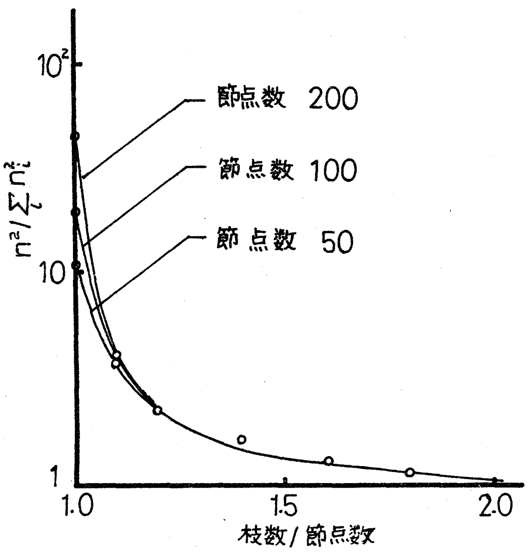

図5. 非可分成分に分割したときの$n^2/\sum_i n_i^2$と枝数/節点数の関係

5.3 連結ランダムグラフの非可分成分数

連結ランダムグラフでは，大部分の節点が，1つの非可分成分に含まれていることが推測される。本節では，この推測を実証するために，連結ランダムグラフに含まれる非可分成分の個数，ならびに3本以上の枝を持つ非可分成分の個数について測定を行う。図6，7にその結果を示す。これより，ほとんどの連結ランダムグラフは，3本以上の枝を含む非可分成分を1つしか持たないことがわかる。

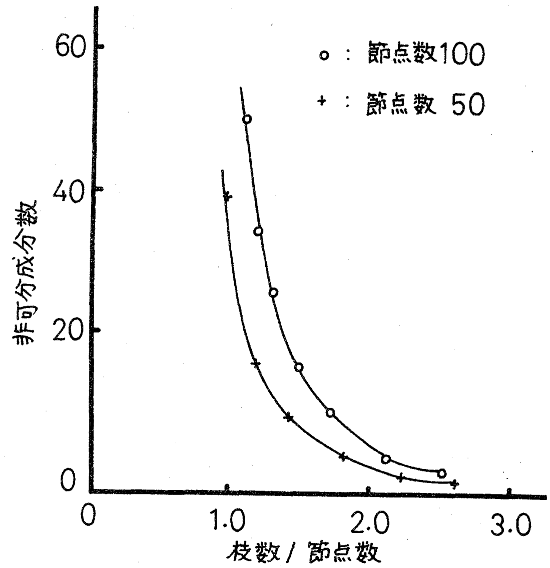

図6．非可分成分の個数と枝数/節点数の関係

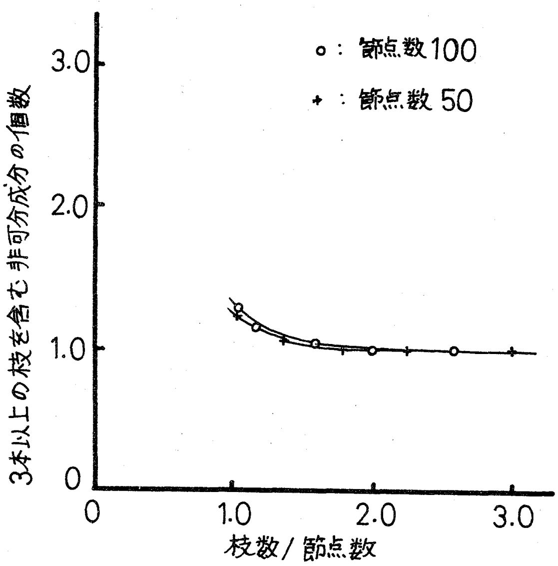

図7．3本以上の枝を含む非可分成分の個数と枝数/節点数の関係

5.4 有向ランダムグラフの強連結成分数

前節までで明らかになったランダムグラフの特徴は，有向ランダムグラフについても当てはまることが予想される。ここでは，有向弱連結ランダムグラフの強連結成分数，ならび

に，節点を2つ以上持つ強連成分の個数について測定を行う。図8,9は，その結果を示したものである。これより，有向ランダムグラフの場合にも，大部分の節点が1つの強連結成分に含まれる特徴があることがわかる。

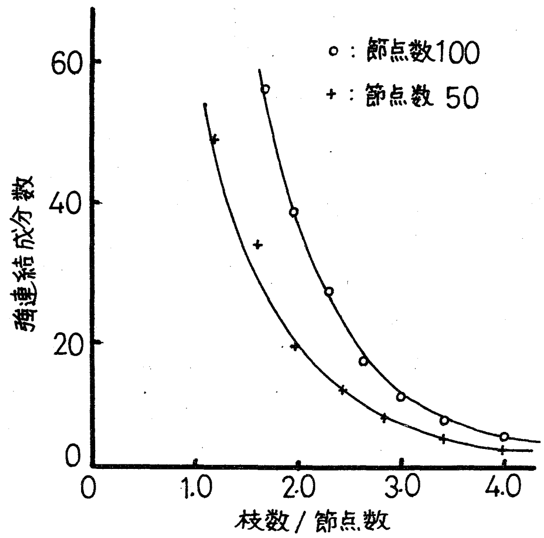

図8. 強連結成分の個数と枝数／節点数の関係

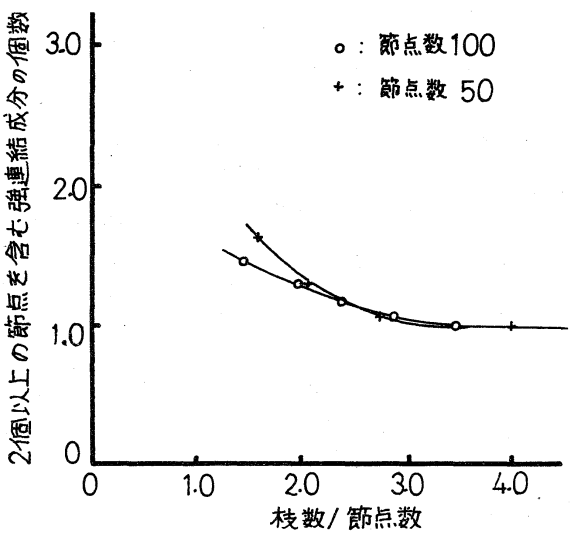

図9. 2個以上の節点を含む強連結成分の個数と枝数／節点数の関係

6. あとがき

筆者らが開発したグラフ処理プログラム－GRAMP－を用いて，ランダムグラフの統計的解析を行った。本文で取り上げたランダムグラフでは，連結成分の大部分が孤立節点であることが示されていたが，本文では，これを確認するとともに，連結ランダムグラフの非可分成分ならびに有向ランダムグラフの強連結成分についても，同様な特徴があることを実験的に明らかにした。

このような結果をみる限り，本文で扱ったランダムグラフと，我々がランダムな回路図などから想起するランダムグラフとの間には大きな隔たりがあるように思われる。ランダムグラフを用いた計算機実験を行う場合には，このことを充分考慮する必要があろう。今後は，この点を考慮した新しいランダムグラフの定義等の研究が待たれる。

本研究の一部は，文部省科学研究費補助金：総合研究A135017(昭51)「ネットワーク構造を持つシステムに関する研究」ならびに，奨励研究A175174(昭51)「大規模システム及びネットワークの計算機処理に関するグラフ理論的研究」の援助のもとに行われた。

参考文献

(1) M. R. Garey and D. S. Johnson: "The complexity of near-optimal graph coloring", J. ACM, Vol. 23, 1, pp. 43-49 (Jan. 1976).

(2) 西関，小川，斎藤："高次極大カットについて"，信学会回路とシステム研資 CST76-95 (1976-10).

(3) 滝内，高見沢，西関，斎藤："グラフ処理言語 -GRAMP-"，信学会回路とシステム研資 CST76-117 (1976-12).

(4) P. Erdos and A. Renyi: "On random graphs I. ", Publicat. Math. Debrecen, 6, pp. 290-297 (1959).

(5) P. Erdos and A. Renyi: "On the existance of a factor of degree one of a connected random graph", Acta Math. Acad. Sci. Hungaricae, 17, pp. 359-368 (1966).

(6) 高見沢，滝内，西関，斎藤："ランダムグラフの統計解析" 信学会回路とシステム研資 CST76-122 (1976-12).

(7) D. E. Knuth: "The Art of Computer Programming, Vol. 2/Seminumerical Algorithm", Addison-Wesley, Reading (1969).